Convergence of random products of Markov transition functions

By S. R. FOGUEL in Jerusalem (Israel)

1. Introduction

The purpose of this note is to apply a theorem of AMEMIYA and ANDO [1] to study a Markov process $\{x_n\}$ with the phase space (X, Σ) which is not stationary but *there exists a mapping* $n \to r(n)$ *of the set of all non-negative integers onto the finite set* $\{1, \ldots, N\}$ *such that the conditional probabilities*

$$\Pr\{x_{n+1} \in A | x_n = x\} \qquad (x \in X,\ A \in \Sigma,\ n = 0, 1, 2, \ldots)$$

depend (besides on A and X) only on $r(n)$, *i.e.*

$$\Pr\{x_{n+1} \in A | x_n = x\} = P_{r(n)}(x, A).$$

By means of $P_j(x, A)$ $(j = 1, \ldots, N)$ one defines in the Banach space of the finite measures on (X, Σ) operators $\nu \to \nu P_j$ by

$$(1.1) \qquad (\nu P_j)(A) = \int P_j(x, A)\, \nu(dx).$$

Thus if μ is a probability measure on (X, Σ) and we choose it to be the distribution of x_0 then the distribution μ_n of x_n is, as well known,

$$\mu_n = \mu\, P_{r(0)} \cdots P_{r(n-1)}.$$

We shall prove the convergence of $\mu_n(A)$ as $n \to \infty$ under suitable conditions.

2. Convergence of random products

Let the mapping $n \to r(n)$ be defined on the set of all non-negative integers and assume the values $1, 2, \ldots, N$. Let $P_j(x, A)$ be subtransition functions on the measure space (X, Σ, λ) with λ being a σ-finite subinvariant measure: $P_j(x, A)$ is a function on $X \times \Sigma$ which is for each $x \in X$ a non-negative measure of total measure $\leqq 1$ and, for each $A \in \Sigma$, a measurable function and

$$(2.1) \qquad \int P_j(x, A)\, \lambda(dx) \leqq \lambda(A).$$

It is well known that $P_j(x, A)$ induce contraction operators on $L_2(X, \Sigma, \lambda)$ by

$$(P_jf)(x) = \int P_j(x, dy)f(y) \tag{2.2}$$

From now on P_j will be considered as contraction operators on L_2 only, also every relation between functions will be a. e.

In [1] it is proved that if each P_j satisfies the condition:

$$\|P_jf\| = \|f\| \quad \text{implies } P_jf=f \tag{W'}$$

and each $1 \leqq j \leqq N$ appears infinitely often in the sequence $r(n)$ then the sequences of operators $P_{r(n)} \dots P_{r(1)}$, $P_{r(1)} \dots P_{r(n)}$ both converge weakly to the projection on the intersection $\bigcap_{j=1}^{N} \{f\colon P_jf=f\}$. Let us study property (W′) in our case.

Lemma 1. *Let $P(x, A)$ be a subtransition function on (X, Σ, λ) with λ being a subinvariant measure. Let $K=\{f\colon \|Pf\|=\|f\|\}$. Then K is generated by characteristic functions of sets of finite measure and if $1_A \in K$[1] for some $A \in \Sigma$ then $P(x, A)$ assumes the values zero and one only.*

Proof. The equation $\|Pf\| = \|f\|$ is equivalent to $P^*Pf=f$ since P is a contraction operator. Thus K is a subspace of L_2.

Now if $f \geqq 0$ then $Pf \geqq 0$. Also if $f \geqq 0$ then $P^*f \geqq 0$: otherwise, if $P^*f<0$ on a set A of positive finite measure, then $0 > \int_A P^*f\,d\lambda = \int P1_A \cdot f\,d\lambda \geqq 0$. Thus if $f \in K$ and is real, $P^*P|f| \geqq |P^*Pf| = |f|$. Inequality is impossible since $\|P^*P\| \leqq 1$, hence $|f| \in K$. Now if $0 \leqq f \leqq c$ then $0 \leqq Pf \leqq c$, also $P^*f \leqq c$: otherwise, if $P^*f > c$ on a set of positive finite measure A, then

$$c\lambda(A) < \int_A P^*f\,d\lambda = \int fP1_A\,d\lambda = \int f(x)P(x, A)\lambda(dx) \leqq c\int P(x, A)\lambda(dx) \leqq c\lambda(A).$$

Therefore if $0 \leqq f \in K$ and $c>0$ then $P^*P(\min(f, c)) \leqq f$ and $P^*P(\min(f, c)) \leqq c$ hence $P^*P(f-\min(f, c)) = f - P^*P(\min(f, c)) \geqq f - \min(f, c)$. Inequality is impossible, hence $f - \min(f, c) \in K$ and therefore $\min(f, c) \in K$.

Thus the conditions of [2], Lemma 1, are satisfied and K is generated by the characteristic functions it contains. Let us conclude the proof by showing that if $1_A \in K$ then $P1_A = P(x, A)$ is a characteristic function:

Put $f = P1_A$ then $0 \leqq f \leqq 1$ and $P^*f = 1_A$.

---

[1] $1_A(x) = \begin{cases} 1 & \text{for } x \in A, \\ 0 & \text{for } x \in X \setminus A. \end{cases}$

Now if $B_\varepsilon=\{x: f(x)>\varepsilon\}$ and C is a set of finite λ measure disjoint to A then $0\leqq\int P^*1_{B_\varepsilon}\cdot 1_C\,d\lambda\leqq\frac{1}{\varepsilon}\int P^*f\cdot 1_C\,d\lambda=0$ or $P^*1_{B_\varepsilon}\leqq 1_A=P^*f$. Thus

$$0\geqq\int P^*(1_{B_\varepsilon}-f)\cdot 1_A\,d\lambda=\int(1_{B_\varepsilon}-f)\cdot f\,d\lambda\to\int(1_B-f)\cdot f\,d\lambda$$

where $B=\{x: f(x)>0\}$. Now $1_B\geqq f$ hence $f=1_B$.

Lemma 2. *Let $P(x, A)$ satisfy the conditions of Lemma* 1. *The corresponding operator P on L_2 satisfies* (W′) *provided:*

(2. 3) $\qquad P(\cdot, A)=1_B$ *and* $\lambda(A)<\infty$ *implies* $1_B=1_A$.

Proof. Condition 2. 3 implies that every characteristic function in K is left invariant by P thus K itself is invariant under P by Lemma 1.

Remark. If $P(\cdot, A)=1_B$ this means that whenever there is a positive probability to move from x to A then the process moves x to A surely. Let us call such a set A a "trap set". Thus condition (2. 3) can be rephrased:

Every "trap set" captures only its own members.

Theorem. *Let $r(n)$ assume the values* 1, ..., *N infinitely often. Let μ be a probability measure which is absolutely continuous with respect to λ and put $d\mu=f\,d\lambda$, $0\leqq f\in L_1(X, \Sigma, \lambda)$. If $P_j(x, A)$ $(1\leqq j\leqq N)$ satisfy* (2. 3) *then for every set A with $\lambda(A)<\infty$*

$$\lim_{n\to\infty}(\mu P_{r(0)}\dots P_{r(n)})(A)=\mu_0(A)$$

where $d\mu_0=f_0\,d\lambda$ and f_0 is the conditional expectation of f on the field $\Sigma'=\bigcap_{j=1}^{N}\{B: \lambda(B)<\infty, P_j(\cdot, B)=1_B\}$.

Proof. With no loss of generality we may assume that $f\in L_2(X, \Sigma)$. It is well known that the conditional expectation f_0 of f on the field Σ' is equal to the orthogonal projection $P'f$ of f on the subspace L' of $L_2(X, \Sigma, \lambda)$, spanned by Σ'.

By the Amemiya—Ando theorem $P_{r(0)}\dots P_{r(n)}$ and therefore $P^*_{r(n)}\dots P^*_{r(0)}$ converge weakly to the orthogonal projection P' on L'. Thus

$$(\mu P_{r(0)}\dots P_{r(n)})A=\int(P_{r(0)}\dots P_{r(n)}1_A)\,d\mu=$$
$$=\int(P_{r(0)}\dots P_{r(n)}1_A)f\,d\lambda=\int 1_A P^*_{r(n)}\dots P^*_{r(0)}f\,d\lambda$$

tends to $\int\limits_A P'f\,d\lambda=\int\limits_A f_0\,d\lambda=\mu_0(A)$.

Let us consider the case where the only "trap sets", for the subtransition functions $P_j(x, A)$, are the trivial sets. Then if λ is not finite, $K=\{0\}$ and $\lim(\mu P_{r(1)}\dots P_{r(n)})(A)=0$ whenever $\lambda(A)<\infty$. On the other hand if $\lambda(X)<\infty$ then $K=\{\text{constant functions}\}$ and $\lim\limits_{n\to\infty}(\mu P_{r(1)}\dots P_{r(n)})(A)=\lambda(A)\,\mu(X)/\lambda(X)$.

References

[1] I. AMEMIYA and T. ANDO, Convergence of random products of contractions in Hilbert space, *Acta Sci. Math.*, **26** (1965), 239–244.

[2] S. R. FOGUEL, On order preserving contractions, *Israel J. Math.*, **1** (1963), 54—59.

(Received May 31, 1966)